Advanced Simulation & Computing™

Cluster Architecture: Standardized Dystopia or Teething Pains?

Ron Brightwell

Scalable Computing Systems

Sandia National Laboratories

Albuquerque, New Mexico, USA

2004 Workshop on Communication Architectures for Clusters (CAC)

April 26, 2004

National Nuclear Security Administration

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Sandia National Laboratories

Advanced Simulation & Computing™

Everybody knows collective ops are a problem, so that's sure to be fixed soon…

- Collective operations can be a system problem
  - See ASCI Q "rogue OS" effects
  - See Barney's talk
- Solution must work with MPI
  - Don't want to end up with IP broadcast semantics
- Probably will need asynchronous collective ops

- If offload is good for collective operations, why isn't it good for peer operations?

NNSA National Nuclear Security Administration

Sandia National Laboratories

Advanced Simulation & Computing™

Are we in danger of standardizing cluster dystopia – if so, how do we fix it?

- Yes, we are in danger
- Possible fixes
  - Change programming model from MPI to something that better fits RDMA semantics
    - Huge cost
  - Change RDMA semantics to something that better fits MPI
    - Smaller cost
  - Change expectations
    - Latency – µs/dollar
    - Bandwidth – MB/s/dollar
    - Overhead – CPU percentage/dollar

NNSA National Nuclear Security Administration

Sandia National Laboratories

Advanced Simulation & Computing™

Existing Performance Bottlenecks*

- Network CPU utilization limits CPU-bound application
  - Any size message
    - Protocol overhead (interrupts, ack processing, etc.)
  - Medium/long messages
    - Receive copy overhead
  - Short messages
    - Kernel bypass (completions are the issue)
- Bandwidth limits
  - Receive copy limits single stream BW to the bcopy rate of a single CPU

*Slide from “The Case for RDMA” by Jim Pinkerton

NNSA National Nuclear Security Administration

Sandia National Laboratories

Advanced Simulation & Computing™

MPI Over RDMA

- Requires CPU involvement in every data transfer to perform MPI matching
- Any size message
  - Protocol overhead (credit-based flow control, acks)
- Medium/long messages
  - Rendezvous protocol performance determined by
    - OS – process must be scheduled
    - Application – must make MPI library calls
- Short messages
  - Receiving process must poll memory

NNSA National Nuclear Security Administration

Sandia National Laboratories

Advanced Simulation & Computing™

Is there really something wrong with this picture, or is the discomfort misplaced?

- **There really is something wrong**

NNSA National Nuclear Security Administration

Sandia National Laboratories

Advanced Simulation & Computing™

Do we "just" need to implement the standards "right", with lower overhead, higher bandwidth, etc., and all will be well?

- Probably not
- Need better understanding of applications
  - See my previous talk ☺
- I'm not bored -- I'm clueless
  - I don't even have enough data to understand the impact of offload, overlap, independent progress
  - I don't know what happens at scale
    - Scale will be more important if network prices continue to drop

NNSA National Nuclear Security Administration

Sandia National Laboratories

Advanced Simulation & Computing™

Do we "just" need better tools? If so, what are they? How do they differ from today's tools?

- **Better benchmarks**
  - **Strong correlation with applications**
- **Better instrumentation**
  - **Measure what is important**
- **Better analysis**
  - **Help figure out what is important**
- **Better development platforms and simulators**
- **Better application build environments**
- **Better application developers ☺**

NNSA National Nuclear Security Administration

Sandia National Laboratories

Advanced Simulation & Computing™

Do we need new programming or system-architectural paradigms? If so, why?

- Better programming model maybe
- Partitioned global address space languages have potential
  - Decrease the complexity of the transport layer
  - Decrease the complexity of network resource management
- But also have many drawbacks
  - Increased compiler complexity
  - Strict SPMD model
  - Don’t support libraries well

NNSA National Nuclear Security Administration

Sandia National Laboratories

Advanced Simulation & Computing™

Why do we have such a problem with commercial software (Grid Services, J2EE, .Net, etc)?

- We care about performance and scalability
- Application developers haven't made us care yet

NNSA National Nuclear Security Administration

Sandia National Laboratories

Advanced Simulation & Computing™

Do we simply accept that clusters just do not work well except for problems exhibiting large-grain embarrassing parallelism?

- No
- We figure out what we think the problems are and publish papers about them ☺

NNSA National Nuclear Security Administration

Sandia National Laboratories

Advanced Simulation & Computing™

Have low cost and peak numbers caused use of clusters where they just don't fit?

- Yes, to some extent
- Low cost may be a problem
  - Buggy hardware
    - End-to-end reliability
  - Insufficient support for MPI
    - MPI works over sockets
  - Not driven by HPC needs
    - HPC isn’t easy

NNSA National Nuclear Security Administration

Sandia National Laboratories

My Pet Peeve

- **Need to do more research and analysis when things work then when they don't**